PERSPECTIVES

INNOVATION

Molecular imaging by mass spectrometry — looking beyond classical histology

Kristina Schwamborn and Richard M. Caprioli

Abstract | Imaging mass spectrometry (IMS) using matrix-assisted laser desorption ionization (MALDI) is a new and effective tool for molecular studies of complex biological samples such as tissue sections. As histological features remain intact throughout the analysis of a section, distribution maps of multiple analytes can be correlated with histological and clinical features. Spatial molecular arrangements can be assessed without the need for target-specific reagents, allowing the discovery of diagnostic and prognostic markers of different cancer types and enabling the determination of effective therapies.

Mass spectrometry (MS) has been the focus of technology development and application for imaging for several decades, with early work in secondary ion mass spectrometry (SIMS) reported in the 1960s[1,2]. SIMS uses an atomic or molecular cluster beam to irradiate a sample with concomitant desorption of ions, which are subsequently analysed to measure their mass to charge ratio (m/z)[3]. This was later followed by laser desorption methods that used a focused laser to achieve the desorption process. One of the limitations of these techniques is that molecules with a molecular mass of more than approximately 2,000 Da cannot be desorbed intact, and so applications typically address the analysis of elements, atomic clusters, drugs and other low molecular mass compounds within tissue sections[4–6]. Matrix-assisted laser desorption ionization (MALDI)-based imaging mass spectrometry (IMS) was introduced in 1997 and provided the capability of measuring both low and high molecular mass compounds, such as proteins with masses of greater than 50,000 Da[7]. In general, the practical image resolution for the desorption technologies for the measurement of intact molecules is around 1 µm, and this is mostly limited by sensitivity.

Much of the early work in MALDI IMS was focused on proteomics in which proteins of molecular masses from 2 to 80 kDa or higher can be ablated and directly analysed intact from tissue sections. Their relative abundance and spatial distribution throughout the section can be measured with a lateral resolution of up to 10 µm using current technology[8,9]. Formalin-fixed and paraffin-embedded (FFPE) samples, constituting most of the tissue collected and stored by pathologists worldwide, can also be analysed by MALDI IMS by combining commonly used antigen retrieval techniques, such as heat-induced antigen retrieval coupled with tryptic digestion of proteins *in situ*[10,11]. In contrast to more commonly used imaging modalities, such as immunohistochemistry (IHC), MALDI IMS allows for the unbiased analysis of tissue sections for the discovery and identification of proteins that are linked to specific tissue types or disease phenotypes as it does not require target-specific reagents[12]. The spatial distribution of hundreds of proteins can be evaluated in parallel and visualized from the same section[13]. By leaving molecular distributions intact, MALDI IMS enables the discovery of spatial molecular arrangements in disease to help assess, for example, the aggressiveness of the disease in discrete areas in the tissue and ultimately to substantially enable the prediction of patient outcome.

MALDI IMS is becoming widely accepted, as shown by the rapidly increasing number of publications in recent years. This interest is driven by the need to gain a better and more fundamental understanding of molecular events involving proteins as well as other types of biological compounds in the development and progression of disease, especially cancer. As the proteome is more complex and dynamic than the genome, it is an effective sensor for ongoing molecular alterations[14,15]. MALDI IMS technology has the capability to overcome some of the limitations of other approaches in the identification of new biomarker molecules that are involved in the development and progression of cancer. Its major strength lies in its ability to measure molecules with direct correlation to anatomical features in tissues at the cellular level, bypassing cumbersome and time-consuming preparation techniques such as laser capture microdissection[16]. This is of particular interest for the analysis of small histological features that would require many sections to be microdissected to acquire the necessary number of cells for proteomic analysis. Another key feature of MALDI IMS is its ability to discover molecular signatures of disease; these signatures typically comprise 5–20 or more different proteins that together result in robust diagnostic patterns[17]. Several studies have shown the usefulness and advantages of this technology in the field of cancer research as an aid to diagnosis or for ascertaining prognosis of different cancer types, as well as in the determination of the effectiveness of a therapeutic regime. This Perspective article highlights some recent publications using MALDI IMS in the field of cancer research, but it is not intended as a complete review of the field.

Technology

MS is recognized as a powerful tool for analysing various different analytes, including small molecules, lipids, DNA segments, peptides and proteins[18–24]. MALDI IMS is amenable to the analysis of high molecular mass biologicals of 100 kDa or more, as well as small molecules with a molecular mass of

less than 1 kDa. For analysis by MALDI, a sample is mixed or coated with an energy-absorbing matrix and subsequently irradiated with a laser beam. In positive ionization mode, singly protonated molecular ions ($[M+H]^+$) are generated from analytes in the sample, whereas in negative ionization mode singly deprotonated ions ($[M-H]^-$) are generated. These are subsequently detected by one of several different types of mass analysers. Time-of-flight (TOF) analysers are commonly used in MALDI IMS to measure the ionized analytes. Following acceleration at a fixed potential[25], ions are separated and recorded based on their mass (molecular mass) to charge ratio (*m/z*) (FIG. 1). As the charge for MALDI is typically one, the analyser read-out measures the protonated or deprotonated molecular mass of the molecule.

The introduction of MALDI IMS first enabled the analysis of intact biological tissue sections in either an ordered array of spots to obtain an image or at selected positions of interest (known as histology directed profiling)[7,26]. The major advantages of these approaches are that they enable the direct correlation of mass spectra with anatomical or pathological features as tissue sections remain intact throughout the analysis and can be stained after MS analysis if necessary[27,28]. Combining well-established antigen retrieval techniques and on-tissue tryptic digestion allows FFPE tissue samples to be analysed by MALDI IMS analysis through the measurement of the resulting peptides[11]. FFPE samples constitute the vast majority of samples that are available in worldwide tissue banks in which millions of samples exist. This type of fixation has also become the standard practice in pathology — samples are processed and embedded in a standardized manner and can be easily stored at room temperature. Moreover, samples from rare diseases are available in addition to the corresponding clinical information[29].

Sample handling and preparation are crucial for achieving high quality and reproducibility in MALDI IMS studies[30,31]. Pre-analytical steps should be standardized and carried out with great care to achieve a high degree of reproducibility. Thin tissue sections can be collected on various sample targets (such as a gold plate or glass slide with conductive surface) and subsequently coated with the matrix. Matrix application can be achieved in several different ways such as dry coating, sublimation, spray coating, and manual and robotic spotting[32–35]. The preferred method of matrix application depends on the analyte. Although dry coating and sublimation are the favoured matrix application techniques for lipid analysis, they are not optimal for protein and peptide analyses; these samples require spray coating or spotting techniques. For these analytes, high quality mass spectra are obtained when the matrix solution crystallizes on the tissue and thereby 'traps' these analytes in the crystals themselves. In FFPE tissue, matrix application is preceded by on-tissue tryptic digestion to release peptide domains that are not cross-linked[36]. A typical workflow for the analysis of fresh frozen tissue is shown in FIG. 2. The acquired data can be visualized by depicting single proteins or peptides (producing protein or peptide images) or by using statistical methods to combine multiple proteins in a model to reveal a signature of key molecular changes within the tissue (producing classification images).

The use of different classification algorithms facilitates data analysis to generate images of tissue sections through two different general approaches: unsupervised (if no knowledge of the tissue composition is available) and supervised (if spatial details of the section are known)[37,38]. Using virtual z-stacks and three-dimensional volume rendering allows the reconstruction of MALDI IMS data to generate three-dimensional volumes, and these can be co-registered with additional data such as block face optical images and magnetic resonance imaging data[39,40].

In summary, MALDI IMS is a versatile technology that can be used for the direct analysis of biological compounds from tissue sections. Although MALDI has been successfully implemented for certain aspects of DNA analysis (such as the analysis of DNA methylation, small oligonucleotides and single nucleotide polymorphisms[22]), analysis of intact DNA from tissue sections has not yet been achieved. MALDI IMS studies have mostly focused on proteins and small molecules in a wide variety of applications, ranging from fundamental biology to clinical studies. Proteomic data obtained from intact tissue sections can be directly correlated with the histology.

Despite impressive advances in MALDI IMS technology, challenges remain to further increase the sensitivity in order to sample deeper into the proteome. The proteins detected are typically at mid- to high-abundance levels. In addition, higher spatial resolution needs to be achieved, ultimately

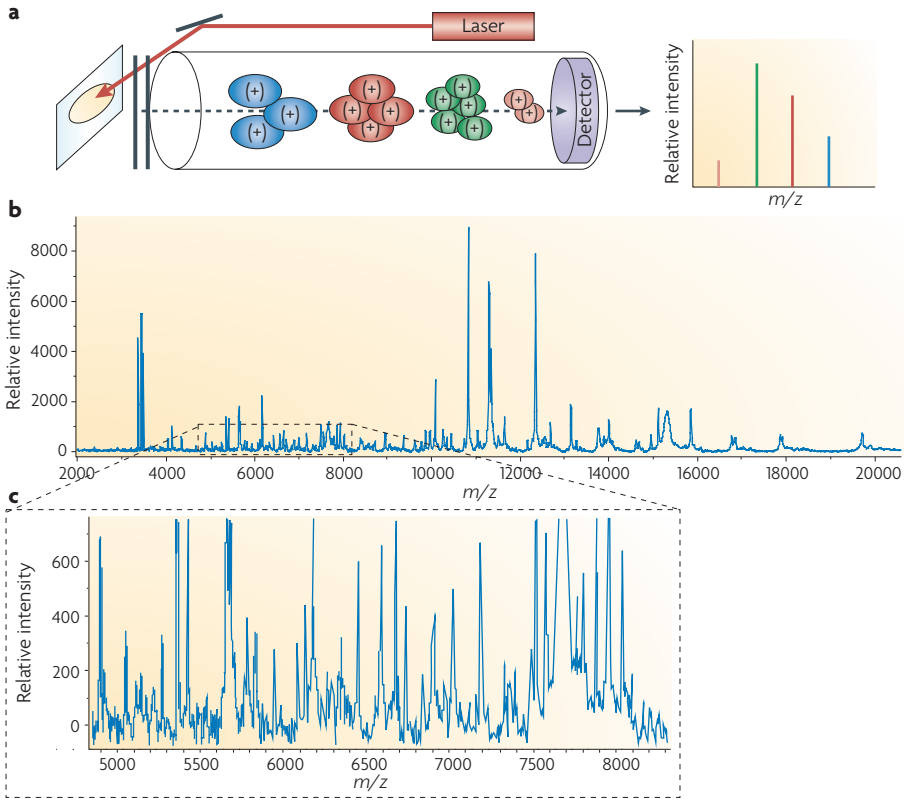

Figure 1 | **Principles of MALDI-TOF IMS. a** | Schematic outline of a typical separation of analytes in a linear matrix-assisted laser desorption ionization (MALDI)-time-of-flight (TOF) imaging mass spectrometer (IMS) based on their mass to charge ratio (*m/z*). **b** | Typical mass spectrum in the mass range between 2 and 20 kDa. **c** | A magnification of the mass range between 5 and 8 kDa.

enabling subcellular analyses. Current limitations in matrix application (such as the size of matrix drops applied to the tissue), as well as laser beam sizes, limit the achievable lateral image resolution to 10–20 μm. Although IHC can attain higher spatial resolution, MALDI IMS has the capability of analysing hundreds of analytes in parallel. A particular strength of MALDI IMS is its independence from the requirement of antibodies. As analytes can be visualized without the need for upstream labelling approaches, discovery becomes an integral advantage of MALDI IMS.

Oncology applications

Applications of MALDI IMS in the field of cancer are diverse and are mainly focused on the analysis of peptides, proteins and small molecules. Some recent studies are discussed below that emphasize specific aspects of MALDI IMS technology. For protein analysis, three different types of studies can be distinguished: diagnostic studies comparing different tissue types (such as tumour versus normal) to aid in pathological diagnosis; prognostic studies to categorize patients with long- or short-term survival; and drug response studies to predict a patient's response to a certain treatment.

MALDI IMS-based studies have been used to elucidate molecular signatures from samples with different tumour types and grades, including brain[41], oral[42], lung[36,43], breast[44,45], gastric[46], pancreatic[47], renal[48], ovarian[49,50] and prostate cancer[28,51]. These studies have been conducted in a retrospective manner and are mainly focused on the identification of molecular signatures of a disease or disease status. As these molecular signatures are comprised of peptides and proteins, identification of the proteins is important, and this can be validated by IHC or western blot. As the technology provides an unbiased method for the analysis of samples as well as multiplexed spatially resolved molecular information, it allows the study of differences in molecular expression levels between adjacent anatomical and pathological structures without the need for labelling approaches, as is the case for IHC. As no a priori knowledge of the proteins is necessary, the discovery of new, unexpected proteins and pathways can be accomplished.

Peptide and protein analysis — diagnostic studies. Studies using MALDI IMS have revealed differences in protein expression between tumour tissue and the immediately adjacent normal tissue[52]. Most interestingly, these differences could still be observed in the surrounding histologically normal tissue,

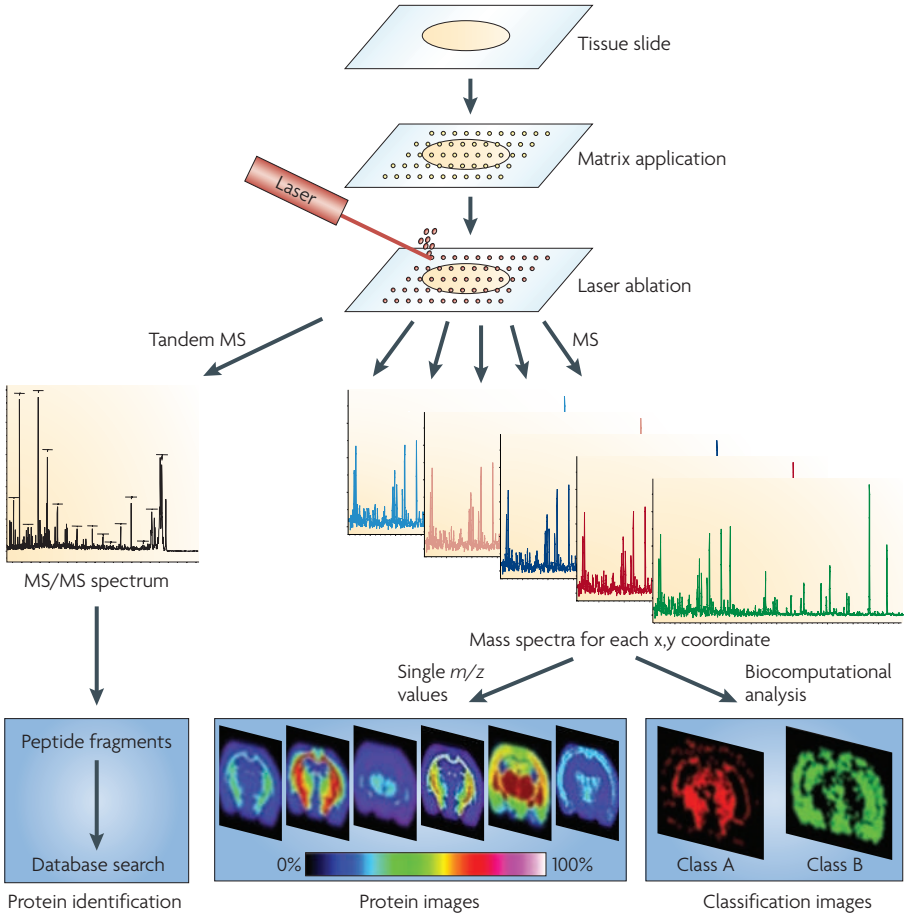

Figure 2 | **Workflow for MALDI IMS analysis.** Schematic outline of a typical workflow for fresh frozen tissue samples. Sample pretreatment steps include cutting and mounting the tissue section on a conductive target. Matrix is applied in an ordered array across the tissue section and mass spectra are generated at each x,y coordinate for protein analysis or tandem MS (MS/MS) spectra for protein identification. Further analytical steps include the visualization of the distribution of a single protein within the tissue (protein image) or statistical analysis to visualize classification images as well as database searching to identify the protein. The scale represents the relative intensity of the protein. IMS, imaging mass spectrometry; MALDI, matrix-assisted laser desorption ionization; MS, mass spectrometry; *m/z*, mass to charge ratio.

decreasing with distance to 1.5 cm beyond the tumour[53]. For example, the molecular assessment of clear cell renal cell carcinoma (ccRCC) samples ($n = 34$), comprising tumour and adjacent normal tissue in the same section, revealed molecular changes in the normal tissue that was adjacent to the tumour that were similar to changes observed in the tumour itself[48]. One group of proteins, which has been shown to be involved in the mitochondrial electron transport system (such as cytochrome *c*; $m/z = 12{,}272$), was consistently underexpressed in tumour and adjacent histologically normal tissue. These findings directly address the issue of the molecular assessment of surgical tumour margins and the desire to reduce local tumour recurrence.

The application of MALDI IMS for diagnostic purposes to differentiate between tumour and normal tissue or different tumour subtypes has become an emerging field as investigators seek to find better markers to aid diagnosis. In a study on prostate cancer, investigators compared a total of 31 fresh frozen prostate cancer and 41 normal prostate biopsy samples to identify differentially expressed peptides[51]. Combining 3 peptides in a genetic algorithm-based model resulted in the correct classification of cancerous areas in 85% of the discovery set (cancer = 11; normal = 10) and 81% of the validation set (cancer = 23; normal = 31). One of the proteins identified was MEKK2 ($m/z = 4{,}355$). These findings were further validated by western blot analysis and IHC, both of which verified that MEKK2 was overexpressed in prostate cancer tissue and cell lines.

In a similar study of ovarian cancer, fresh frozen ovarian cancer tissue samples ($n = 25$) were analysed in comparison to benign

ovarian tumours ($n = 23$)[49]. A fragment of the 11S proteasome activator complex REGα (also known as proteasome activator subunit 1 (PSME1); $m/z = 9{,}744$) was found to be overexpressed in the cancer samples. Results were confirmed by western blot analysis as well as IHC. IHC revealed a distinct and diagnosis-dependent localization in cellular compartments, with cytoplasmic localization of REGα in carcinomas as well as nuclear staining but no cytoplasmic staining in 76.9% of benign tumours. This is the first report describing high expression levels of REGα in ovarian cancer. The 11S regulator complex (PA28; of which REGα is a subunit) binds to the 20S proteasome, and downregulation of PA28 in tumour cells has been shown to result in an impaired presentation of tumour antigens[54].

Two separate studies of non-small-cell lung cancer (NSCLC) were able to identify proteomic patterns to accurately classify and predict histopathological groups (such as adenocarcinoma and squamous cell carcinoma)[36,43]. In the first study, investigators analysed fresh frozen samples from 79 lung tumours and 14 normal lung tissues[43]. Using a training set of 42 tumour samples and 8 normal samples, investigators created a model that could correctly classify all samples in the training set as well as all samples in an independent validation set (37 cancer and 6 normal samples). Moreover, expression profiles of two proteins also allowed the classification of tumours with and without lymph node metastasis with 85% and 75% accuracy in the training and validation sets, respectively. Combining 15 peaks in a proteomic pattern enabled the distinction between patients with NSCLC with poor ($n = 25$) and good ($n = 41$) prognosis ($p<0.0001$). Two proteins ($m/z = 10{,}519$ and $m/z = 4{,}964$) that were highly expressed in primary NSCLC were identified as SUMO2 and thymosin-β4, X-linked (TMSβ4X), respectively.

Another study of NSCLC demonstrated the use of FFPE tissue microarray (TMA) samples for high-throughput analysis and classification by means of on-tissue tryptic digestions and MALDI IMS[36]. Using a support vector machine algorithm-based model enabled the correct classification of all squamous cell carcinoma ($n = 22$) and all adenocarcinoma ($n = 18$) samples. MALDI-tandem mass spectrometry (MS/MS) analysis directly from the TMA section allowed the identification of approximately 50 proteins. For example, heat shock protein β1 (HSPB1) identified by three tryptic peptides, was found to be almost exclusively expressed in squamous cell carcinomas. All three tryptic peptides showed a similar ion density distribution across the TMA, thereby corroborating the derivation of these peptides from the same protein. Two other peptides ($m/z = 1{,}407.7$ and $m/z = 1{,}410.7$) were identified as a fragment of cytokeratin 6 (CK6) and CK5 (also known as keratin 5), respectively. These two antigens are usually expressed in squamous cell carcinoma, whereas adenocarcinomas are mainly CK5- and CK6-negative. Commercially available antibodies stain both antigens. However, a wide variation in the expression of both cytokeratins in patients was revealed when comparing MALDI IMS-based peptide images of CK5 and CK6 with IHC staining using an antibody against CK5 and CK6 (REF. 75) (FIG. 3).

In breast cancer, MALDI IMS has been used to classify ERBB2 (also known as HER2) receptor status on stored fresh frozen tissue samples. The assessment of ERBB2 expression in breast cancer is crucially important for treatment decisions in newly diagnosed patients with primary breast cancer to predict which patients are most likely to respond to trastuzumab therapy[44]. Fresh frozen samples from 48 patients were analysed. The expression of ERBB2 was evaluated by IHC and fluorescence *in situ* hybridization in the training set (ERBB2-positive $n = 15$ and ERBB2-negative $n = 15$) and by IHC only in the validation set (ERBB2-positive $n = 6$ and ERBB2-negative $n = 12$). Classification of the ERBB2 receptor status in the training set using an artificial neuronal network-based model resulted in 87% sensitivity and 93% specificity. Applying the same model to predict the ERBB2 receptor status in the validation set resulted in the correct classification of 16 of the 18 cases (83% sensitivity and 92% specificity). In addition, a protein overexpressed in ERBB2-positive samples was identified as cysteine-rich intestinal protein 1 (CRIP1; $m/z = 8{,}403$). This result confirms previous findings of the co-expression of *CRIP1* and *ERBB2* mRNAs[55].

Peptide and protein analysis — prognostic studies. In a study on patient outcome, samples from 108 patients with glioma and 19 non-cancer patients were subjected to MALDI IMS analysis to obtain protein signatures that correlated with tumour histology and patient survival[41]. In addition to being able to distinguish between non-cancer subjects and patients with glioma with an average of >92% accuracy, investigators could also identify a proteomic signature (24 peaks) to differentiate between patients from two prognostic groups: a short-term survival group (52 patients; mean survival of <15 months) and a long-term survival group (56 patients; mean survival of >90 months). Multivariate analysis showed a strong correlation between the identified protein expression signature and patient survival, demonstrating that the protein expression signature is an independent indicator of patient survival. Additionally, for a subgroup of the patients with glioma — those with glioblastoma ($n = 57$) — two different peaks

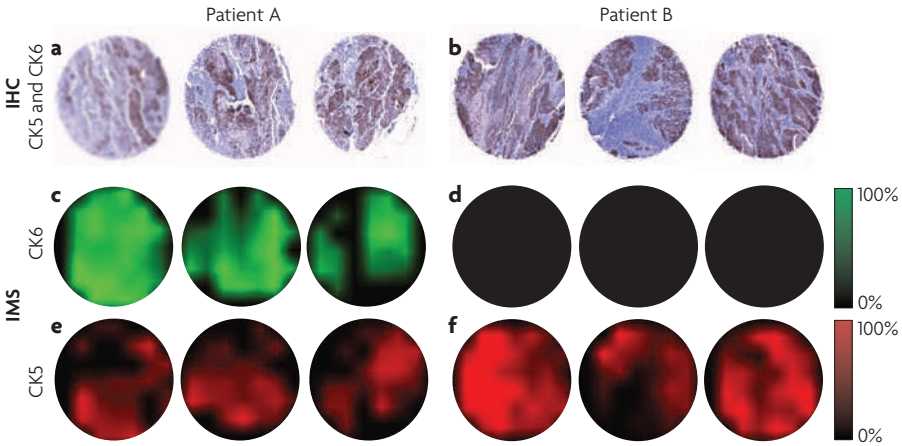

Figure 3 | **Comparison of IHC and MALDI IMS results.** Immunohistochemical (IHC) staining with an antibody against cytokeratin 5 (CK5) and CK6 of triplicate lung cancer biopsy samples from two different patients (patient A or patient B) showing positive staining in all biopsy samples (parts **a** and **b**). The corresponding peptide image of a 1407.7 Da fragment of CK6 (green; parts **c** and **d**) reveals this peptide to be present only in one patient, whereas the 1410.7 Da fragment of CK5 (red; parts **e** and **f**) is present in all biopsy samples[75]. The scale represents the relative intensity of the peptide. IMS, imaging mass spectrometry; MALDI, matrix-assisted laser desorption ionization. Image courtesy of M.R. Groseclose, Vanderbilt University, USA.

could separate patients with short-term survival (28 patients; average survival of 10.9 months) from patients with long-term survival (29 patients; average survival of 16.8 months). This proteomic signature could also be verified as an independent predictor of patient survival by a multivariate Cox proportional hazards model. In total, six proteins could be identified by high-performance liquid chromatography (HPLC) and MALDI-MS/MS in this study: calcyclin (also known as S100A6; m/z = 10,092), dynein light chain LC8-type 2 (DLC8B; also known as DYMLL2; m/z = 10,262), calpactin I light chain (also known as S100A10; m/z = 11,073), phosphoprotein-enriched in astrocytes (PEA15; m/z = 15,035), fatty acid-binding protein 5 (FABP5; m/z = 15,076) and tubulin-folding cofactor A (TBCA; m/z = 17,268). Two of these proteins, S100A6 and DLC8B, were among the classifiers for overall patient prognosis of all patients with glioma, with S100A6 overexpressed in patients with short-term survival and DLC8B predominant in patients with long-term survival. S100A6, a protein thought to be involved in tumorigenesis and cell proliferation, has previously been described as overexpressed in other tumours, such as colon cancer, and its expression level in epithelial cells was found to be directly proportional to the increase in malignancy[56,57]. DLC8B has been reported to interact with BIM, a pro-apoptotic protein, to negatively regulate its apoptotic function[58].

Peptide and protein analysis — drug response studies. Another goal of MALDI IMS-based proteomic studies is the identification of markers that correlate with patient response to a therapeutic regime. MALDI IMS, together with gene expression profiling, has been used to identify markers of taxane sensitivity in patients with breast cancer receiving a neoadjuvant therapeutic treatment regime of the taxane paclitaxel with radiation[45]. Fresh frozen pretreatment tissue samples from 19 patients were subjected to histology-directed profiling. Of these 19 patients, 6 achieved a pathological complete response and 13 showed residual disease or no response after treatment. Comparing spectra from tumour regions between the two groups revealed three highly overexpressed (>30-fold) features (m/z = 3,371, m/z = 3,442 and m/z = 3,485) and four features of lower expression (m/z = 5,667, m/z = 6,955, m/z = 7,007 and m/z = 15,348) in the responder group. The three significantly overexpressed proteins had previously been identified as α-defensins (DEFα1, m/z = 3,371; DEFα2, m/z = 3,442; and DEFα3, m/z = 3,485) in proteomic analyses of different types of cancer, such as head and neck cancer[59]. To verify the tumour cells as the source of the DEFα expression and to eliminate the possibility of an artefact caused by blood contamination of the tissue or the presence of neutrophils, IHC was carried out. Samples from patients with pathological complete responses exhibited a positive staining pattern of DEFα in the tumour cells, whereas samples from non-responders showed little or no staining of DEFα.

In another study it was shown that early proteomic changes assessed by MALDI IMS could predict the treatment response of mammary tumours in transgenic mice[60]. Fresh frozen tumour sections from mouse mammary tumour virus (MMTV);*Erbb2*-transgenic mice treated with different doses of a small molecule tyrosine kinase inhibitor of epidermal growth factor receptor (EGFR), erlotinib, and/or an ERBB2-blocking antibody (trastuzumab), were analysed by MALDI IMS at various time points and compared with tumour sections from untreated mice. Tumours from mice treated with erlotinib showed a dose- and time-dependent decrease in TMSβ4X (m/z = 4,965), as well as ubiquitin (m/z = 8,565) and an increase of a fragment of E-cadherin-binding protein E7 (also known as c-CBL-like protein 1 (CBLL1); m/z = 4,794) when compared with tumours from untreated mice. Those drug-induced changes precede the inhibition of proliferation or cell death, as measured by IHC (using the detection of proliferating cell nuclear antigen (PCNA)) and terminal deoxynucleotidyl transferase (Tdt)-mediated nick end labelling (TUNEL). Additional proteins found to be more highly expressed in tumours treated with both drugs include a fragment of calmodulin, acyl-CoA binding protein (also known as DBI), calgizzarin (also known as S100A11), histone H3 and histone H4.

Lipid analysis. Interest in assessing lipid changes in disease has increased considerably owing to their important functions in signal transduction and energy storage[61]. Altered levels of lipids are found in many human diseases, including cancer[62]. MALDI IMS for spatially resolved analysis of lipids has also grown rapidly as the technology has developed. Owing to their low molecular mass (<1,000 Da), *in situ* analysis can be hampered by interference from other components in the tissue and through the normal tissue preservation protocols used in histology[61]. Initial MALDI IMS studies using fresh frozen tissue sections suggest that regional differences in concentrations for specific molecular species of lipids can be detected in tissues[63,64]. Investigators were able to analyse the distribution of mitochondrion-specific lipids, the cardiolipins, in different rat organ sections[65]. Work from our laboratory in studies of biopsy samples from patients with ccRCC showed marked differences in lipid distributions between tumour and normal regions (unpublished observations, S. Puolitaival and R.M.C.). In total, 78 matched pairs were analysed in positive and negative ionization mode. Most linoleic acid-containing phospholipids were found to be more abundant in tumour regions, whereas sphingomyelins were more abundant in normal regions. Although these findings are interesting in terms of understanding the total molecular changes that occur in tumorigenesis, their specific role in this process remains unclear. In a recently published non-cancer-related study, MALDI IMS was used to successfully characterize the spatial and temporal distribution of phospholipid species during mouse embryo implantation[66].

Applications of MALDI IMS to drug discovery. The assessment of the distribution and metabolism of drug candidates in targeted tissues and throughout the body is of central interest to drug discovery and in the understanding of drug responses[67,68]. Drug analysis uses MS/MS that allows both the molecular species and the structure-specific fragments to be simultaneously monitored to increase sensitivity and better validate the identification of signals with high confidence. Single reaction monitoring (SRM) and multiple reaction monitoring (MRM) techniques further enable this analysis to be accomplished in a high-throughput manner[69]. These techniques monitor the structure-related composition of the drug (precursor or parent ion) in the mass spectrometer to form one (SRM) or multiple (MRM) specific fragment ions.

Minimizing the tissue handling and washing steps that are typically a part of sample pretreatment protocols for the analysis of proteins is crucial for the detection of small molecules, as such treatment can degrade or compromise the level of the drug of interest. Initial studies by MALDI IMS have shown the possibility of detecting not only the distribution of the drug itself but also the simultaneous distribution of its individual metabolites in whole-body tissue sections[70]. FIGURE 4 shows an example of small

molecule imaging in whole-animal sections in which the orally administered anti-cancer drug is present in the digestive system as well as the tumour (personal communication, M. L. Reyzer, R. A. Smith and H. C. Manning).

A recently published study used MALDI IMS to detect the localization of the alkylating agent oxaliplatin and its derivates in heated intra-operative chemotherapy-like treated rat kidneys[71]. The lowest concentration of oxaliplatin detectable by MALDI IMS was determined as 0.23 ± 0.05 mg of total oxaliplatin per gram of tissue using a 15 µm- thick tissue section. Imaging results showed that oxaliplatin and its derivates could be detected within the kidney cortex only. Penetration of the drug into the medulla was not observed. Another study imaged the distribution of vinblastine within rat whole-body sections in order to visualize the distribution of the drug in different organs[72]. Ion density maps of vinblastine (precursor ion; $m/z = 811.4$) and its fragments (for example, $m/z = 793$) showed a higher distribution in the liver, kidney and tissue surrounding the gastrointestinal tract. However, especially in the kidney, differences between the distribution of the precursor ion together with its fragment at $m/z = 751$ were observed when compared with the other seven fragments, with the first showing a relatively high signal intensity in the renal pelvis. The distribution of orally administered erlotinib and its metabolites was investigated in tissue sections from rat liver, spleen and muscle by MALDI IMS[67]. The highest drug concentrations were found in liver sections together with the detection of one of the major drug metabolites. In addition, direct quantitative analysis of the drug on tissue (liver, spleen and muscle) by MALDI IMS was carried out and compared with standard LC–MS/MS on tissue homogenates. The ratio of total ion intensities from liver and spleen, as estimated by MALDI IMS analysis, was in good agreement with the ratio calculated by LC–MS/MS analysis. One study of tumour response to drug treatment also analysed the spatial localization of elortinib directly in tissue sections by SRM of the transition $m/z = 394.2$ to $m/z = 278.1$ (REF. 60). The drug was predominantly found to be distributed in the more vascularized peripheral areas of the tumour section 16 hours after drug administration.

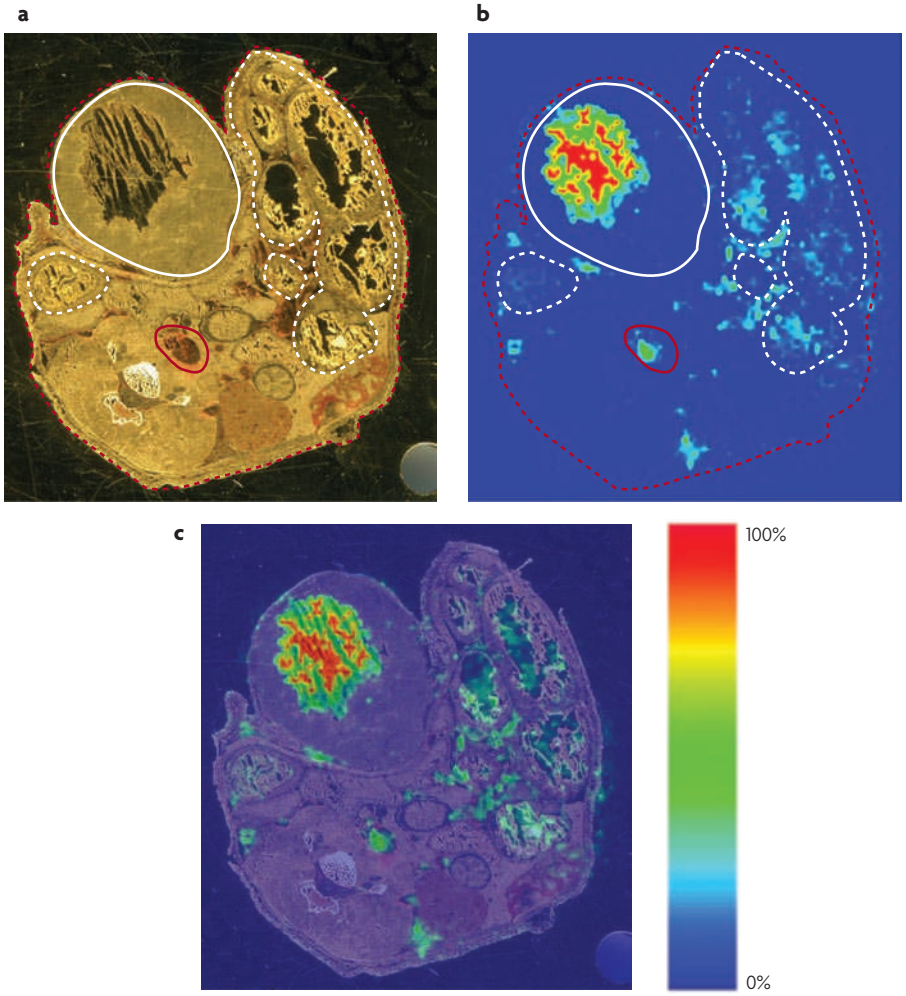

Figure 4 | **Small molecule imaging.** Unstained transverse whole mouse section (part **a**), image of the drug distribution (part **b**) and overlay (part **c**). The orally administered anti-cancer drug is present in high concentrations in the tumour and can also be found in lower concentrations in the digestive system and blood vessels. The solid white line represents the tumour; dashed white line the gastrointestinal tract; solid red line the major blood vessels and the dashed red line the outline of the whole section. The scale represents the relative intensity of the drug. Personal communication, M. L. Reyzer, R. A. Smith and H. C. Manning, Vanderbilt University, USA.

Conclusions and perspectives

MALDI IMS is a new technology with respect to defined clinical applications. Its outstanding molecular recognition capabilities should substantially benefit molecular pathology as the technology develops further. Such development would include the implementation of biocomputational tools as well as the building and validation of extensive databases that correlate molecular signatures with histological and clinical data such as patient outcome. Many studies have shown the strength of this emerging technology and its ability to be used to aid in the diagnosis and prognosis of several different cancer types. In addition to discovering potential new drug targets, MALDI IMS also offers the possibility of analysing the distribution and effect of drug candidates directly in tissues and whole animals. The most important characteristics of MALDI IMS in a discovery setting are its multiplexicity and its independence from target-specific reagents such as antibodies. Further improvements in the technology will bring an increased ability to measure proteins of lower expression levels in tissue sections. Also, as image resolution increases, smaller laser spot sizes (that are subcellular, for example) will require even greater advances in sensitivity as the amount of material analysed per spot will decrease. Progress in many areas has already been shown; for example, in the development of special sample pretreatment protocols to allow the detection of hydrophobic and membrane-bound proteins[73,74].

As this technology has proved its feasibility and versatility, the next phase in its clinical use will be its translation to effective applications in the clinic. More comprehensive studies with many hundreds of samples are needed to validate the robustness and potential use in a clinical setting in support of personalized medicine. The technology holds

great promise for establishing protein or lipid signatures as robust aids in disease diagnosis and most importantly for prediction of a patient's response to therapy and overall prognosis. MALDI IMS, perhaps in conjunction with other approaches, can clearly bring the art and practice of molecular pathology to a new and more effective level.

Kristina Schwamborn and Richard M. Caprioli are at the Mass Spectrometry Research Center, Department of Biochemistry, Vanderbilt University, Nashville, Tennessee 37232-2195, USA.
Correspondence to R.M.C.
e-mail: richard.m.caprioli@vanderbilt.edu

doi:10.1038/nrc2917
Published online 19 August 2010

1. Liebl, H. Ion microprobe mass analyzer. *J. Appl. Phys.* **38**, 5277–5283 (1967).
2. Castaing, R. & Slodzian, G. Microanalysis by secondary ionic emission. *J. Microsc.* **1**, 395–410 (1962).
3. Esquenazi, E., Yang, Y. L., Watrous, J., Gerwick, W. H. & Dorrestein, P. C. Imaging mass spectrometry of natural products. *Nat. Prod. Rep.* **26**, 1521–1534 (2009).
4. Fragu, P., Klijanienko, J., Gandia, D., Halpern, S. & Armand, J. P. Quantitative mapping of 4′-iododeoxyrubicin in metastatic squamous cellcarcinoma by secondary ion mass spectrometry (SIMS) microscopy. *Cancer Res.* **52**, 974–977 (1992).
5. Pacholski, M. L. & Winograd, N. Imaging with mass spectrometry. *Chem. Rev.* **99**, 2977–3006 (1999).
6. Chandra, S. & Lorey, D. R. SIMS ion microscopy in cancer research: single cell isotopic imaging for chemical composition, cytotoxicity and cell cycle recognition. *Cell. Mol. Biol. (Noisy-le-grand)* **47**, 503–518 (2001).
7. Caprioli, R. M., Farmer, T. B. & Gile, J. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal. Chem.* **69**, 4751–4760 (1997).
8. Seeley, E. H. & Caprioli, R. M. Molecular imaging of proteins in tissues by mass spectrometry. *Proc. Natl Acad. Sci. USA* **105**, 18126–18131 (2008).
9. Chaurand, P., Schriver, K. E. & Caprioli, R. M. Instrument design and characterization for high resolution MALDI-MS imaging of tissue sections. *J. Mass Spectrom.* **42**, 476–489 (2007).
10. Lemaire, R. *et al.* Direct analysis and MALDI imaging of formalin-fixed, paraffin-embedded tissue sections. *J. Proteome Res.* **6**, 1295–1305 (2007).
11. Ronci, M. *et al.* Protein unlocking procedures of formalin-fixed paraffin-embedded tissues: application to MALDI-TOF imaging MS investigations. *Proteomics* **8**, 3702–3714 (2008).
12. Zimmerman, T. A., Monroe, E. B., Tucker, K. R., Rubakhin, S. S. & Sweedler, J. V. Chapter 13: imaging of cells and tissues with mass spectrometry: adding chemical information to imaging. *Methods Cell Biol.* **89**, 361–390 (2008).
13. Chaurand, P., Sanders, M. E., Jensen, R. A. & Caprioli, R. M. Proteomics in diagnostic pathology: profiling and imaging proteins directly in tissue sections. *Am. J. Pathol.* **165**, 1057–1068 (2004).
14. Celis, J. E. & Gromov, P. Proteomics in translational cancer research: toward an integrated approach. *Cancer Cell* **3**, 9–15 (2003).
15. Hanash, S. Disease proteomics. *Nature* **422**, 226–232 (2003).
16. Emmert-Buck, M. R. *et al.* Laser capture microdissection. *Science* **274**, 998–1001 (1996).
17. Caprioli, R. M. Deciphering protein molecular signatures in cancer tissues to aid in diagnosis, prognosis, and therapy. *Cancer Res.* **65**, 10642–10645 (2005).
18. Siuzdak, G. The emergence of mass spectrometry in biochemical research. *Proc. Natl Acad. Sci. USA* **91**, 11290–11297 (1994).
19. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198–207 (2003).
20. Schiller, J. *et al.* Matrix-assisted laser desorption and ionization time-of-flight (MALDI-TOF) mass spectrometry in lipid and phospholipid research. *Prog. Lipid Res.* **43**, 449–488 (2004).
21. Harvey, D. J. *et al.* Comparison of fragmentation modes for the structural determination of complex oligosaccharides ionized by matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **9**, 1556–1561 (1995).
22. Tost, J. & Gut, I. G. DNA analysis by mass spectrometry-past, present and future. *J. Mass Spectrom.* **41**, 981–995 (2006).
23. Karas, M. & Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* **60**, 2299–2301 (1988).
24. Tanaka, K. *et al.* Protein and polymer analyses up to *m/z*100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **2**, 151–153 (1988).
25. Mann, M., Hendrickson, R. C. & Pandey, A. Analysis of proteins and proteomes by mass spectrometry. *Annu. Rev. Biochem.* **70**, 437–473 (2001).
26. Cornett, D. S. *et al.* A novel histology-directed strategy for MALDI-MS tissue profiling that improves throughput and cellular specificity in human breast cancer. *Mol. Cell. Proteomics* **5**, 1975–1983 (2006).
27. Crecelius, A. C. *et al.* Three-dimensional visualization of protein expression in mouse brain structures using imaging mass spectrometry. *J. Am. Soc. Mass Spectrom.* **16**, 1093–1099 (2005).
28. Schwamborn, K. *et al.* Identifying prostate carcinoma by MALDI-imaging. *Int. J. Mol. Med.* **20**, 155–159 (2007).
29. Becker, K. F. *et al.* Quantitative protein analysis from formalin-fixed tissues: implications for translational clinical research and nanoscale molecular diagnosis. *J. Pathol.* **211**, 370–378 (2007).
30. Cornett, D. S., Reyzer, M. L., Chaurand, P. & Caprioli, R. M. MALDI imaging mass spectrometry: molecular snapshots of biochemical systems. *Nature Methods* **4**, 828–833 (2007).
31. Schwartz, S. A., Reyzer, M. L. & Caprioli, R. M. Direct tissue analysis using matrix-assisted laser desorption/ionization mass spectrometry: practical aspects of sample preparation. *J. Mass Spectrom.* **38**, 699–708 (2003).
32. Puolitaival, S. M., Burnum, K. E., Cornett, D. S. & Caprioli, R. M. Solvent-free matrix dry-coating for MALDI imaging of phospholipids. *J. Am. Soc. Mass Spectrom.* **19**, 882–886 (2008).
33. Chaurand, P. *et al.* Integrating histology and imaging mass spectrometry. *Anal. Chem.* **76**, 1145–1155 (2004).
34. Aerni, H. R., Cornett, D. S. & Caprioli, R. M. Automated acoustic matrix deposition for MALDI sample preparation. *Anal. Chem.* **78**, 827–834 (2006).
35. Hankin, J. A., Barkley, R. M. & Murphy, R. C. Sublimation as a method of matrix application for mass spectrometric imaging. *J. Am. Soc. Mass Spectrom.* **18**, 1646–1652 (2007).
36. Groseclose, M. R., Massion, P. P., Chaurand, P. & Caprioli, R. M. High-throughput proteomic analysis of formalin-fixed paraffin-embedded tissue microarrays using MALDI imaging mass spectrometry. *Proteomics* **8**, 3715–3724 (2008).
37. Deininger, S. O., Ebert, M. P., Futterer, A., Gerhard, M. & Rocken, C. MALDI imaging combined with hierarchical clustering as a new tool for the interpretation of complex human cancers. *J. Proteome Res.* **7**, 5230–5236 (2008).
38. Hanselmann, M. *et al.* Toward digital staining using imaging mass spectrometry and random forests. *J. Proteome Res.* **8**, 3558–3567 (2009).
39. Andersson, M., Groseclose, M. R., Deutch, A. Y. & Caprioli, R. M. Imaging mass spectrometry of proteins and peptides: 3D volume reconstruction. *Nature Methods* **5**, 101–108 (2008).
40. Sinha, T. K. *et al.* Integrating spatially resolved three-dimensional MALDI IMS with *in vivo* magnetic resonance imaging. *Nature Methods* **5**, 57–59 (2008).
41. Schwartz, S. A. *et al.* Proteomic-based prognosis of brain tumor patients using direct-tissue matrix-assisted laser desorption ionization mass spectrometry. *Cancer Res.* **65**, 7674–7681 (2005).
42. Patel, S. A. *et al.* Imaging mass spectrometry using chemical inkjet printing reveals differential protein expression in human oral squamous cell carcinoma. *Analyst* **134**, 301–307 (2009).
43. Yanagisawa, K. *et al.* Proteomic patterns of tumour subsets in non-small-cell lung cancer. *Lancet* **362**, 433–439 (2003).
44. Rauser, S. *et al.* Classification of HER2 receptor status in breast cancer tissues by MALDI imaging mass spectrometry. *J. Proteome Res.* **9**, 1854–1863 (2010).
45. Bauer, J. A. *et al.* Identification of markers of taxane sensitivity using proteomic and genomic analyses of breast tumors from patients receiving neoadjuvant paclitaxel and radiation. *Clin. Cancer Res.* **16**, 681–690 (2010).
46. Morita, Y. *et al.* Imaging mass spectrometry of gastric carcinoma in formalin-fixed paraffin-embedded tissue microarray. *Cancer Sci.* **101**, 267–273 (2010).
47. Djidja, M. C. *et al.* MALDI-ion mobility separation-mass spectrometry imaging of glucose-regulated protein 78 kDa (Grp78) in human formalin-fixed, paraffin-embedded pancreatic adenocarcinoma tissue sections. *J. Proteome Res.* **8**, 4876–4884 (2009).
48. Oppenheimer, S. R., Mi, D., Sanders, M. & Caprioli, R. M. A molecular analysis of tumor margins by MALDI mass spectrometry in renal carcinoma. *J. Proteome Res.* **9**, 2182–2190 (2010).
49. Lemaire, R. *et al.* Specific MALDI imaging and profiling for biomarker hunting and validation: fragment of the 11S proteasome activator complex, reg α fragment, is a new potential ovary cancer biomarker. *J. Proteome Res.* **6**, 4127–4134 (2007).
50. Kang, S. *et al.* Molecular proteomics imaging of tumor interfaces by mass spectrometry. *J. Proteome Res.* **9**, 1157–1164 (2010).
51. Cazares, L. H. *et al.* Imaging mass spectrometry of a specific fragment of mitogen-activated protein kinase/extracellular signal-regulated kinase kinase kinase 2 discriminates cancer from uninvolved prostate tissue. *Clin. Cancer Res.* **15**, 5541–5551 (2009).
52. Herring, K. D., Oppenheimer, S. R. & Caprioli, R. M. Direct tissue analysis by matrix-assisted laser desorption ionization mass spectrometry: application to kidney biology. *Semin. Nephrol.* **27**, 597–608 (2007).
53. Caldwell, R. L., Gonzalez, A., Oppenheimer, S. R., Schwartz, H. S. & Caprioli, R. M. Molecular assessment of the tumor protein microenvironment using imaging mass spectrometry. *Cancer Genom. Proteom.* **3**, 279–288 (2006).
54. Sijts, A. *et al.* The role of the proteasome activator *PA28* in MHC class I antigen processing. *Mol. Immunol.* **39**, 165–169 (2002).
55. Wilson, K. S., Roberts, H., Leek, R., Harris, A. L. & Geradts, J. Differential gene expression patterns in *HER2*/neu-positive and -negative breast cancer cell lines and tissues. *Am. J. Pathol.* **161**, 1171–1185 (2002).
56. Bronckart, Y. *et al.* Development and progression of malignancy in human colon tissues are correlated with expression of specific Ca^{2+}-binding S100 proteins. *Histol. Histopathol.* **16**, 707–712 (2001).
57. Lesniak, W., Slomnicki, L. P. & Filipek, A. S100A6 - new facts and features. *Biochem. Biophys. Res. Commun.* **390**, 1087–1092 (2009).
58. Puthalakath, H., Huang, D. C., O'Reilly, L. A., King, S. M. & Strasser, A. The proapoptotic activity of the *Bcl-2* family member Bim is regulated by interaction with the dynein motor complex. *Mol. Cell* **3**, 287–296 (1999).
59. Roesch-Ely, M. *et al.* Proteomic analysis reveals successive aberrations in protein expression from healthy mucosa to invasive head and neck cancer. *Oncogene* **26**, 54–64 (2007).
60. Reyzer, M. L. *et al.* Early changes in protein expression detected by mass spectrometry predict tumor response to molecular therapeutics. *Cancer Res.* **64**, 9093–9100 (2004).
61. Jackson, S. N. & Woods, A. S. Direct profiling of tissue lipids by MALDI-TOFMS. *J. Chromatogr B Analyt. Technol. Biomed. Life Sci.* **877**, 2822–2829 (2009).
62. Wenk, M. R. The emerging field of lipidomics. *Nature Rev. Drug Discov.* **4**, 594–610 (2005).
63. Murphy, R. C., Hankin, J. A. & Barkley, R. M. Imaging of lipid species by MALDI mass spectrometry. *J. Lipid Res.* **50**, S317–S322 (2009).
64. Meriaux, C., Franck, J., Wisztorski, M., Salzet, M. & Fournier, I. Liquid ionic matrixes for MALDI mass spectrometry imaging of lipids. *J. Proteomics* **73**, 1204–1218 (2010).
65. Wang, H. Y., Jackson, S. N. & Woods, A. S. Direct MALDI-MS analysis of cardiolipin from rat organs sections. *J. Am. Soc. Mass Spectrom.* **18**, 567–577 (2007).
66. Burnum, K. E. *et al.* Spatial and temporal alterations of phospholipids determined by mass spectrometry during mouse embryo implantation. *J. Lipid Res.* **50**, 2290–2298 (2009).

67. Signor, L. *et al.* Analysis of erlotinib and its metabolites in rat tissue sections by MALDI quadrupole time-of-flight mass spectrometry. *J. Mass Spectrom.* **42**, 900–909 (2007).
68. Cornett, D. S., Frappier, S. L. & Caprioli, R. M. MALDI-FTICR imaging mass spectrometry of drugs and metabolites in tissue. *Anal. Chem.* **80**, 5648–5653 (2008).
69. Kitteringham, N. R., Jenkins, R. E., Lane, C. S., Elliott, V. L. & Park, B. K. Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **877**, 1229–1239 (2009).
70. Khatib-Shahidi, S., Andersson, M., Herman, J. L., Gillespie, T. A. & Caprioli, R. M. Direct molecular analysis of whole-body animal tissue sections by imaging MALDI mass spectrometry. *Anal. Chem.* **78**, 6448–6456 (2006).
71. Bouslimani, A., Bec, N., Glueckmann, M., Hirtz, C. & Larroque, C. Matrix-assisted laser desorption/ionization imaging mass spectrometry of oxaliplatin derivatives in heated intraoperative chemotherapy (HIPEC)-like treated rat kidney. *Rapid Commun. Mass Spectrom.* **24**, 415–421 (2010).
72. Trim, P. J. *et al.* Matrix-assisted laser desorption/ionization-ion mobility separation-mass spectrometry imaging of vinblastine in whole body tissue sections. *Anal. Chem.* **80**, 8628–8634 (2008).
73. Grey, A. C., Chaurand, P., Caprioli, R. M. & Schey, K. L. MALDI imaging mass spectrometry of integral membrane proteins from ocular lens and retinal tissue. *J. Proteome Res.* **8**, 3278–3283 (2009).
74. Norris, J. L., Porter, N. A. & Caprioli, R. M. Combination detergent/MALDI matrix: functional cleavable detergents for mass spectrometry. *Anal. Chem.* **77**, 5036–5040 (2005).
75. Groseclose, M.R. *High-throughput analysis of tissue microarrays of disease: combining in situ proteomics with MALDI imaging mass spectrometry.* Thesis, Vanderbilt Univ. (2009).

Acknowledgements
The authors would like to thank E. H. Seeley for brain images, M. R. Groseclose for lung images, and M. L. Reyzer, H. C. Manning and R. A. Smith for whole mouse images.

Competing interests statement
The authors declare no competing financial interests.

DATABASES
National Cancer Institute Drug Dictionary: http://www.cancer.gov/drugdictionary/erlotinib | oxaliplatin | paclitaxel | trastuzumab | vinblastine
Pathway Interaction Database: http://pid.nci.nih.gov/ERBB2

FURTHER INFORMATION
Richard M. Caprioli's homepage: http://www.mc.vanderbilt.edu/root/vumc.php?site=msrc

ALL LINKS ARE ACTIVE IN THE ONLINE PDF

SCIENCE AND SOCIETY

Integrating biobanks: addressing the practical and ethical issues to deliver a valuable tool for cancer research

R. William G. Watson, Elaine W. Kay and David Smith

Abstract | Cancer is caused by complex interactions between genes, environment and lifestyles. Biobanks of well-annotated human tissues are an important resource for studying the underlying mechanisms of cancer. Although such biobanks exist, their integration to form larger biobanks is now required to provide the diversity of samples that are needed to study the complexity and heterogeneity of cancer. Clear guidelines and policies are also required to address the challenges of integrating individual institutional or national biobanks and build public trust. This Science and Society article highlights some of the main practical and ethical issues that are undergoing discussion in the integration of tissue biobanks for cancer.

It is widely accepted that although basic scientific studies carried out using cell lines and animal models can be informative about the cellular and molecular aspects of cancer there is a clear requirement to confirm this in human samples. The concept of patient-specific and disease-specific ('targeted') therapy has expanded rapidly in recent years. Many researchers believe that this concept of personalized medicine will provide the solution to the considerable challenges posed to the clinical treatment of cancer. The move from the traditional 'one size fits all' approach for the treatment of cancer to targeted approaches seems to offer genuine hope for improved patient outcomes. There are a few examples for which the concept of a highly effective drug treatment targeted towards a specific limited patient population has become reality. The use of imatinib (Gleevec; Norvartis) in chronic myeloid leukaemia[1], the use of monoclonal antibodies that target the epidermal growth factor receptor (EGFR) in patients with EGFR-expressing metastatic colon cancer[2] and the use of the ERBB2 (also known as HER2)-specific monoclonal antibody trastuzumab (Herceptin; Genentech) in ERBB2-positive breast cancers[3] are such examples.

The potential benefits of this personalized approach to the treatment of disease are considerable. They include the identification of improved biological targets using validated biomarker studies, the capacity to increase the likely success of clinical trials by preselecting the patient population and the fact that this will in turn reduce the time, cost and the likelihood of failure of clinical trials[4]. Information from validated biomarker studies also allows the re-introduction of drugs that have failed in a clinical trial setting or that have been withdrawn from the market to be re-applied in a more targeted way. Similarly, biomarker studies might also offer the potential to avoid adverse side effects, and this would, in turn, lead to higher compliance with various treatment regimes.

The need for large integrated biobanks

One of the biggest limiting factors to the successful translation of basic scientific cellular and molecular studies into improved patient outcome has been the lack of access to large, appropriate and well-annotated cohorts of human tissue[5,6]. Focused disease-specific institutional biobanks have had some success in translational and personalized medicine (as described above). However, owing to the complex and heterogeneous nature of cancer, it is now clear that much larger biobanks are required[7,8]. For genetic main-effect studies 2,000–5,000 samples are needed, for lifestyle main-effect studies 2,000–20,000 samples are required and for gene–lifestyle interaction studies 20,000–50,000 samples are required[8]. Only when these larger resources are available can we truly understand the interactions between gene, environment, lifestyle and disease and translate this knowledge into the clinic through innovative diagnostics, therapeutics and preventive strategies for cancer. These larger resources can only be achieved by the integration of existing biobanks that already have a wealth of information and samples. However, there are many obstacles and challenges associated with such integration, including technical, logistical, ethical and legal ones.

Groups across both North America and Europe have started to address these obstacles and challenges to move this process forwards. Initially, national programmes were established that linked previously collected biobank samples. These included the Canadian Tumour Repository Network (CTRNet; see the CTRNet website; Further